Wetland Intrinsic Potential Tool: Identifying forested wetlands through lidar-derived machine learning

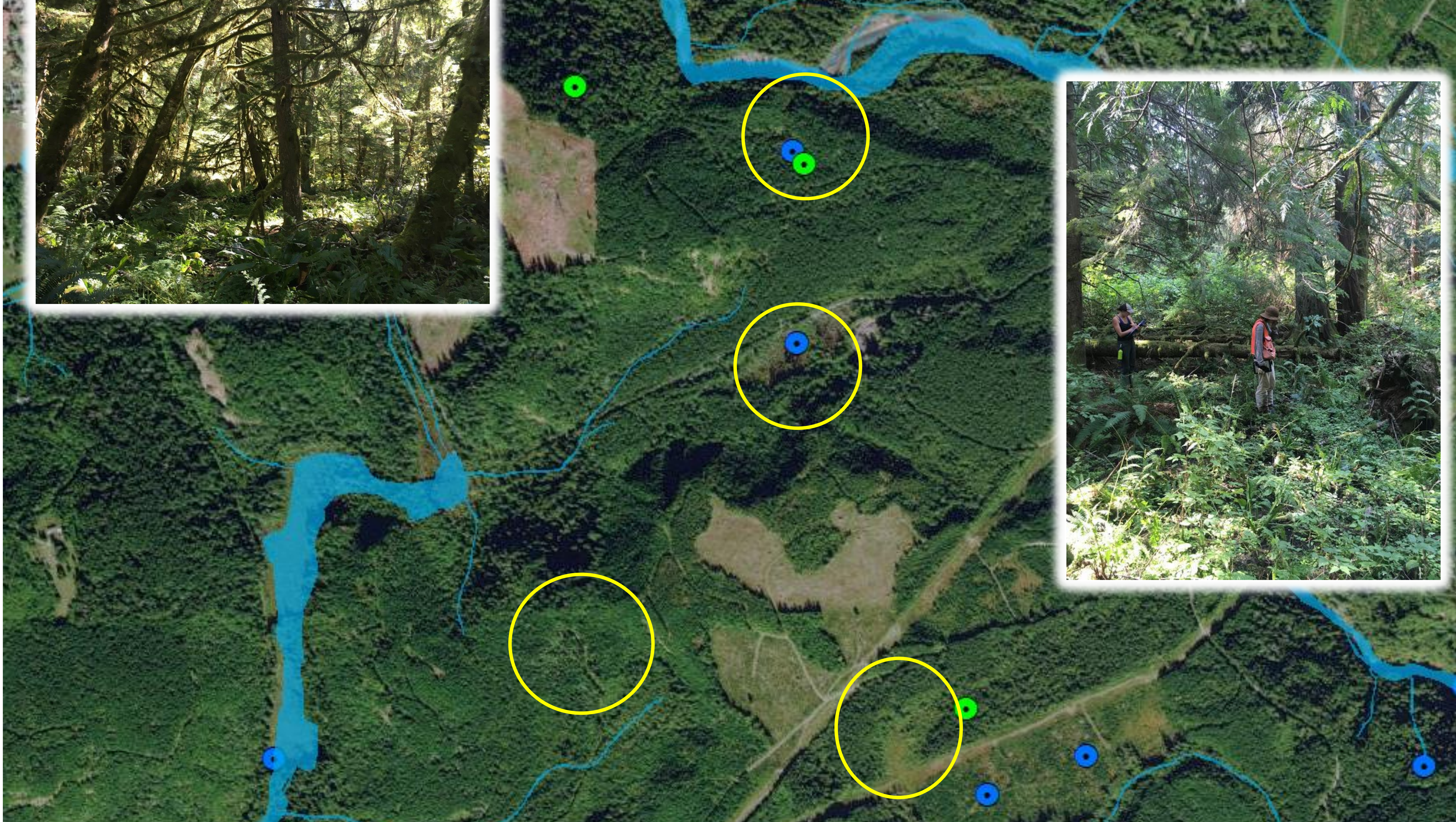
University of Washington
Meghan Halabisky, Monika Moskal, Luke Rogers

TerrainWorks
Daniel Miller, Lee Benda, Kevin Andras

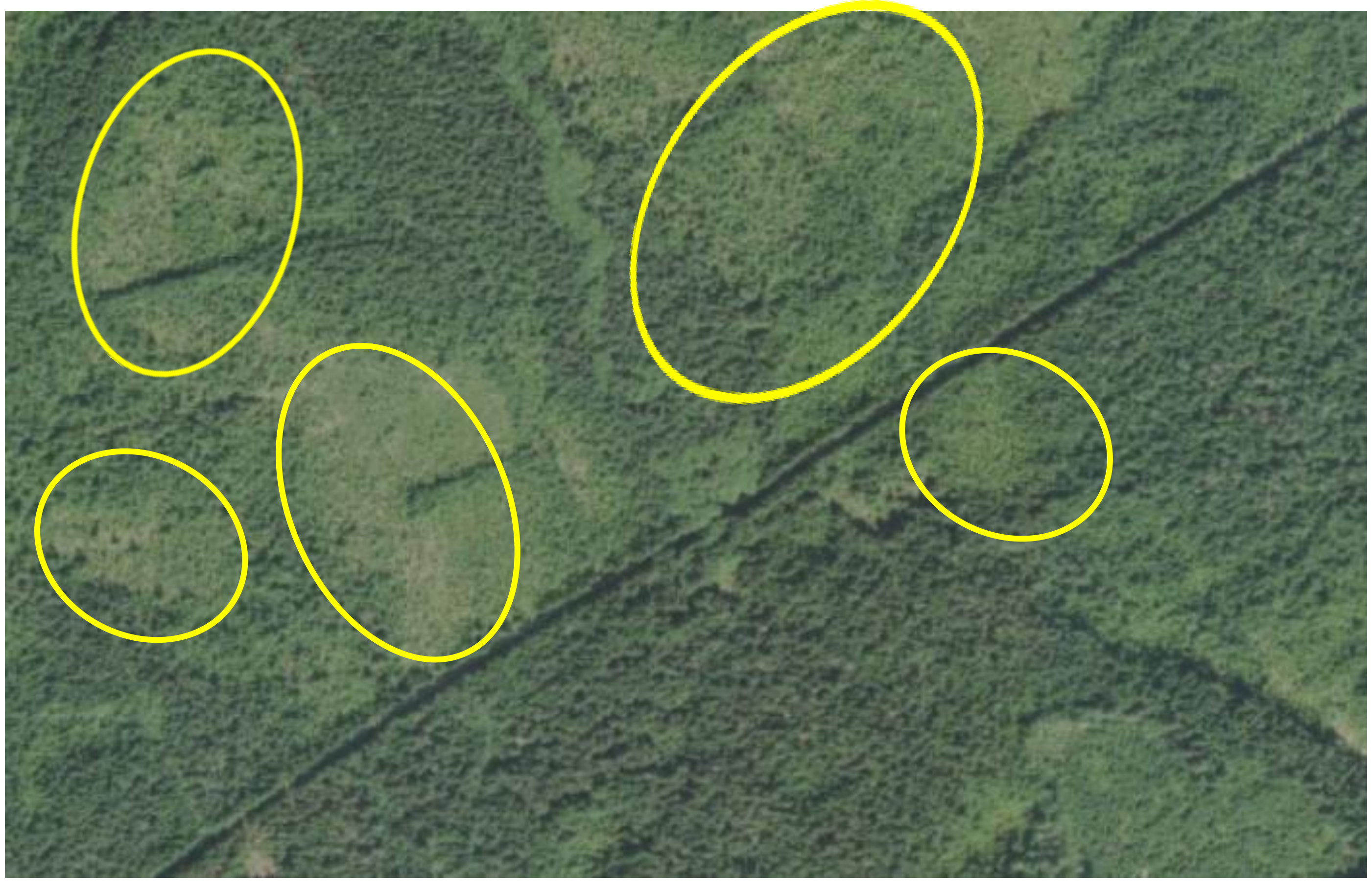
Drone image of Beaver Creek, Mashel Watershed

Problem Statement: Identifying wetlands

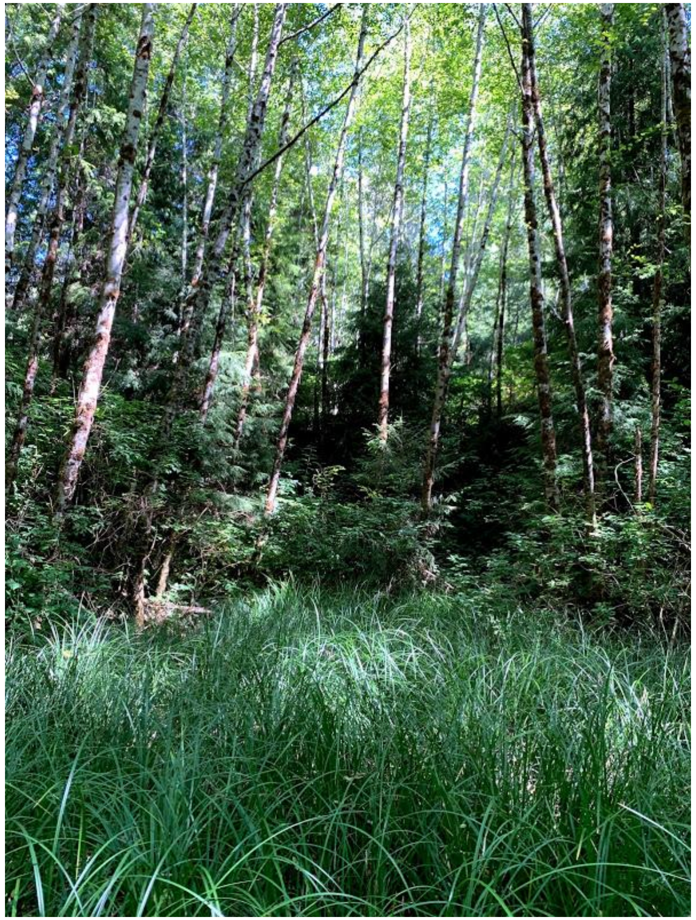
In Washington State available wetland inventories are often out-of-date and have high errors of omission (especially in forested and agricultural areas).



Where are the wetlands on forest lands?

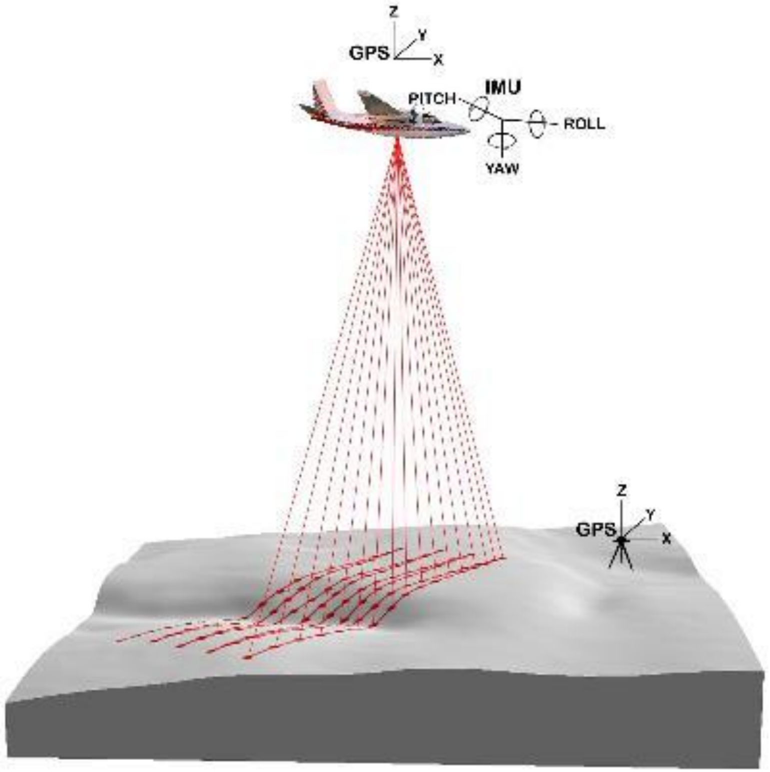


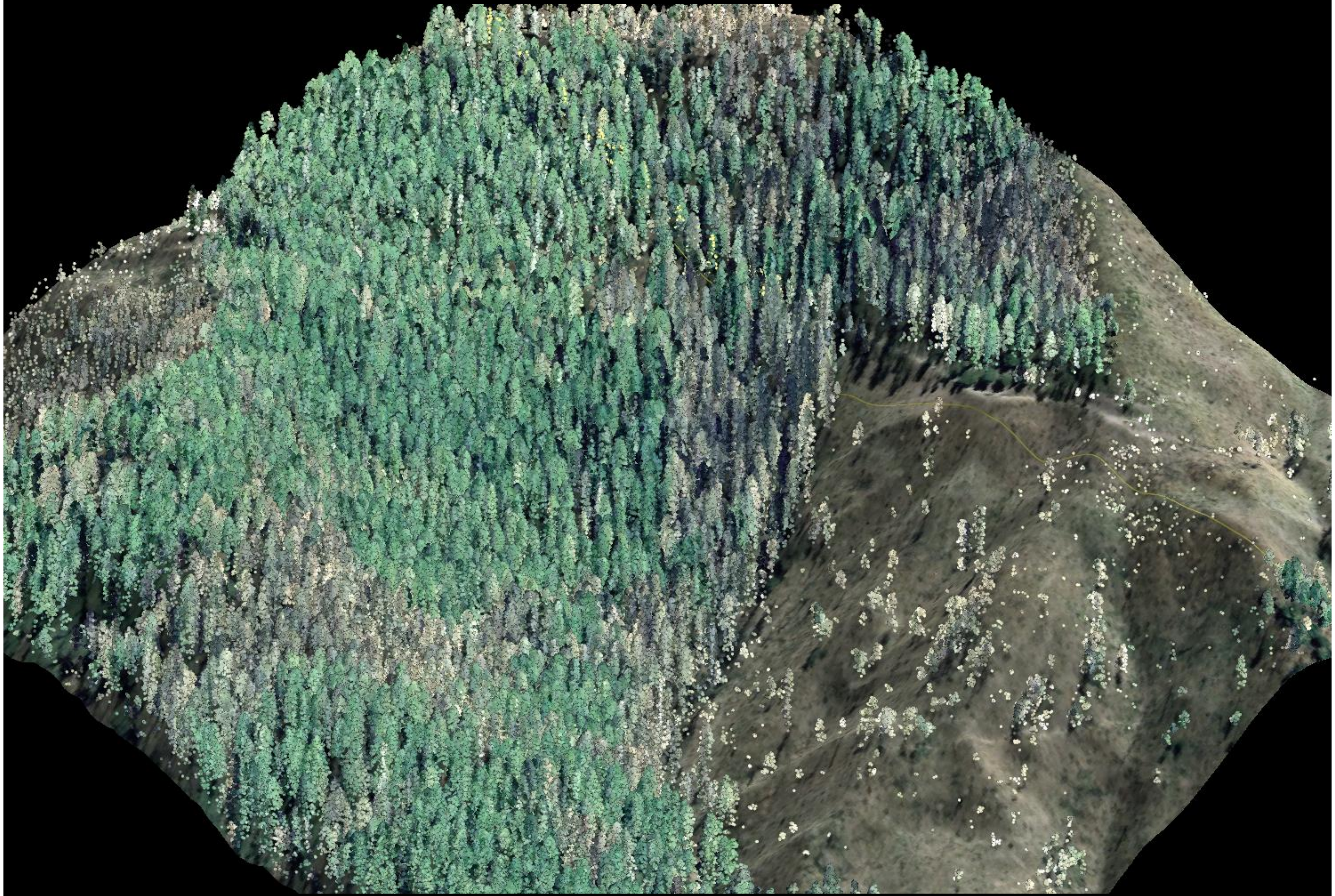
Where are the wetlands on forest lands?

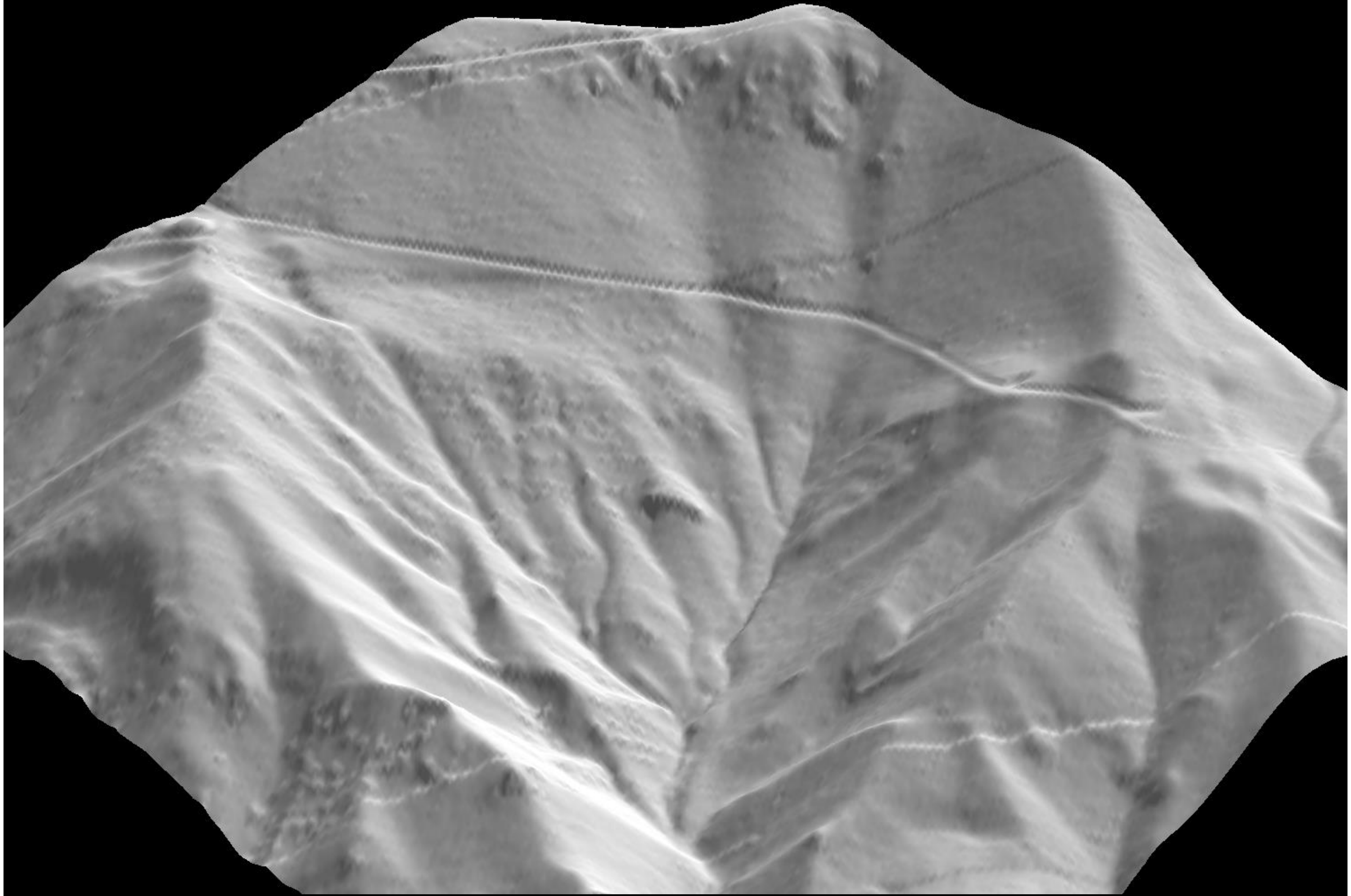


Key technology used: LiDAR - Light Detection and Ranging

- Active airborne laser scanner
- Returns are points with X, Y and Z coordinates
- LiDAR Products:
    - Ground model
    - Canopy Surface model
    - Slope
    - Intensity image







Phase 1: Develop a tool for mapping hydrological and geomorphological controls on wetland occurrence.

The ‘Wetlands Intrinsic Potential’ (WIP) tool uses digital elevation models (LiDAR) and may incorporate other digital data, including soils, geology, and multi-spectral imagery. (Luke Rogers (UW), TerrainWorks)

Phase 2: Use field data on wetland locations to evaluate methods developed in Phase 1, and to develop new machine learning models mapping probability of wetland occurrence. (Meghan Halabisky, TerrainWorks).

WIP Tool Phase 2: Project objectives

1.) Identify key variables used to predict wetlands in the PNW

2.) Collect sample training of wetland and non-wetland locations

4.) Test machine learning methods – random forest models

5.) Develop an ArcGIS tool that is flexible and can be used by anyone as screening tool

METHODS

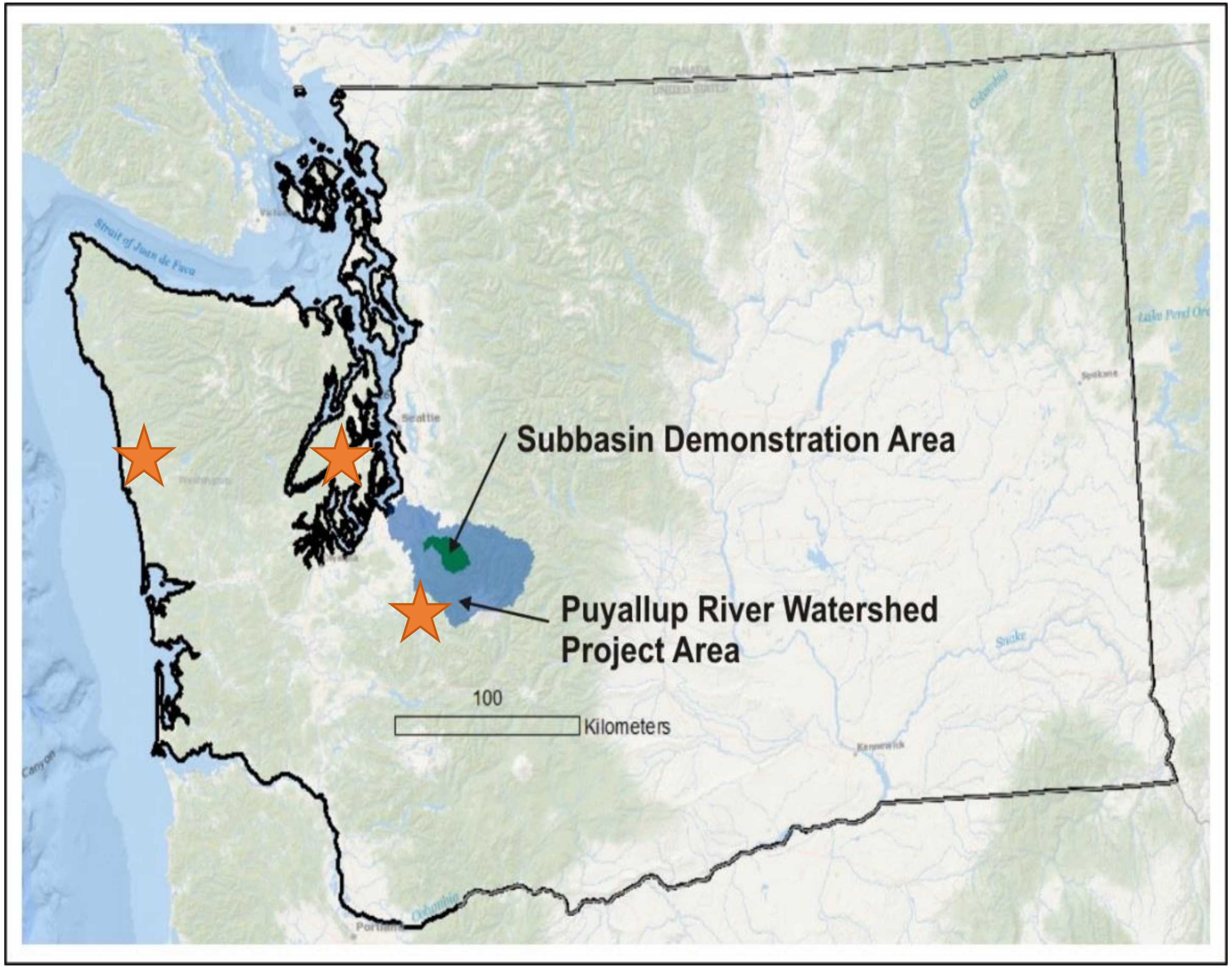
Study Areas

Model development:

- Puyallup watershed

Tested transferability of model on:

- Mashel watershed
- Coulter Creek-Kitsap peninsula
- Hoh watershed



Identify Key Variables

Literature Review:

East coast, Midwest, & E. WA have had success mapping wetlands:

- Topographic wetness index
- Lidar intensity
- Leaf-off imagery
- Depth-to-water index
- Rule based approach v. random forest method

Wetland Mapping in the Upper Midwest United States: An Object-Based Approach Integrating Lidar and Imagery Data

Lian P. Rampi, Joseph F. Knight, and Keith C. Pelletier

Abstract

This study investigated the effectiveness of using high resolution data to map wetlands in three ecoregions in Minnesota. High resolution data included multispectral leaf-off aerial imagery and lidar elevation data. These data were integrated using an Object-Based Image Analysis (OBIA) approach. Results for each study area were compared against field and image interpreted reference data using error matrices, accuracy estimates, and the kappa statistic. Producer's and user's accuracies were in the range of 92 to 96 percent and 91 to 96 percent, respectively, and overall accuracies ranged from 96-98 percent for wetlands larger than 0.20 ha (0.5 acres). The results of this study may allow for increased accuracy of mapping wetlands efforts over traditional remote sensing methods.

Introduction

Wetlands are naturally dynamic systems of important value to the environment and society. The US Army Corps of Engineers (USACE) in cooperation with the US Environmental Protection Agency (EPA) have defined wetlands, incorporating technical and policy considerations, as "…those areas that are inundated or saturated by surface or ground water at a frequency and duration to support and under normal circumstances do support, a prevalence of vegetation typically adapted for life in saturated soil conditions" (Federal Register, 1980 and 1982). Wetlands can reduce some of the negative effects of flooding and recharge groundwater by gradually releasing flood water and snow melt. Wetlands offer habitat that supports wildlife and fishing activities. Wetlands also provide ecosystem services, including educational, aesthetic, and economic opportunities. For example, intact freshwater marshes in Canada have a total economic value of approximately 5,800 USD per hectare compared to 2,400 USD when those lands are drained and used for agriculture (Millennium Ecosystem Assessment, 2005; Turner *et al.*, 2000).

Due to wetland loss and degradation, many of the preceding benefits have been reduced and are increasingly impacted. About 215 million acres of wetlands existed in the United States at the time of European settlement. However, by the mid-1970s, only 99 million acres of the original wetlands remained. Many of the lost wetlands were drained and are currently used for agriculture, resource extraction, urbanization, and other commercial purposes (Dahl and Johnson, 1991; Frayer *et al.*, 1983; Stedman and Dahl, 2008). Minnesota is not an exception to this large national wetland loss. Nearly half of Minnesota's original wetlands were lost due to extensive agricultural drainage and urban development. According to the Minnesota Pollution Control Agency (MPCA) (2006), many original natural wetlands were changed into local storm-water ponds to make additional land available for urban development.

Currently in Minnesota only a few cities have updated wetland inventories. For the rest of Minnesota the only wetland inventory available is the National Wetlands Inventory (NWI). The Minnesota NWI maps were completed in the late 1980s using aerial photos (some black and white) collected between 1979 and 1988 (LMIC, 2007). Several 7.5' quadrangles in northwestern Minnesota and a much larger area in northeastern Minnesota were mapped based on 1970s 1:80 000 scale black-and-white photos (MPCA, 2006). Changes in the landscape have occurred which limit the use of the NWI maps due to the outdated data and techniques used to create them. Thus, there is a need to update wetland inventories with accurate boundaries and improved delineation of smaller wetlands. An updated wetland inventory would provide information that could be used to make accurate decisions for the conservation, protection, and restoration of wetlands. Although a Minnesota statewide update is underway, it is a heavily image interpretation-based project that is not expected to be completed until 2020. Thus, more automated techniques may be useful in the near term.

A fast and effective method to identify accurate wetland boundaries involves the use of remote sensing data and techniques (Butera, 1983; Corcoran *et al.*, 2011; Knight *et al.*, 2013). To the present time, the majority of wetland mapping efforts using remote sensing data and techniques has been focused on evaluating traditional pixel-methods with medium to coarse resolution data. In many cases, the use of remote sensing for wetland mapping has resulted in low accuracy estimates, often due to mixed pixels and insufficient spectral resolution (Grenier *et al.*, 2007; Fournier *et al.*, 2007; Lunetta and Balogh, 1999; Ozesmi and Bauer, 2002). Integration of high resolution optical and elevation data has been shown to reduce the mixed pixel problem (Frohn *et al.*, 2009; Maxa and Bolstad, 2009). Some studies have integrated optical and elevation data to map wetlands using traditional pixel-based methods. However, their accuracy results were low for wetland classification due to the use of low to medium spatial resolution data and pixel-based techniques (Baker *et al.*, 2006; Ozesmi and Bauer, 2002).

An object-based approach may be a better option to integrate high resolution data and overcome some limitations, including the mixed pixel problem and salt-and-pepper effect of traditional pixel-based techniques (Myint *et al.*, 2011; Zhou

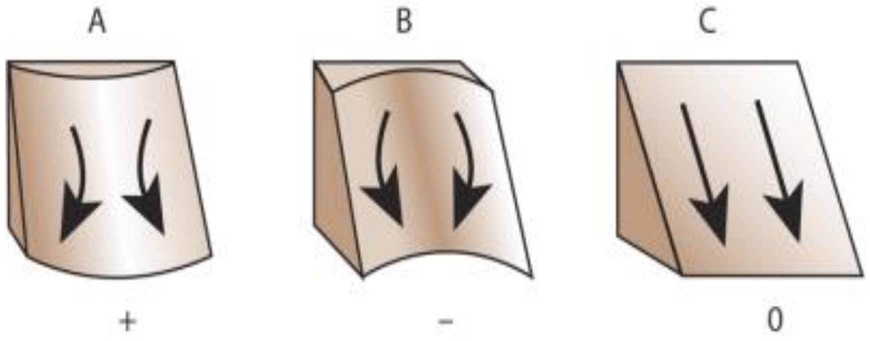
---

The Department of Forest Resources, University of Minnesota, 1530 Cleveland Ave, N., Saint Paul, MN 55108 (ortiz073@umn.edu).

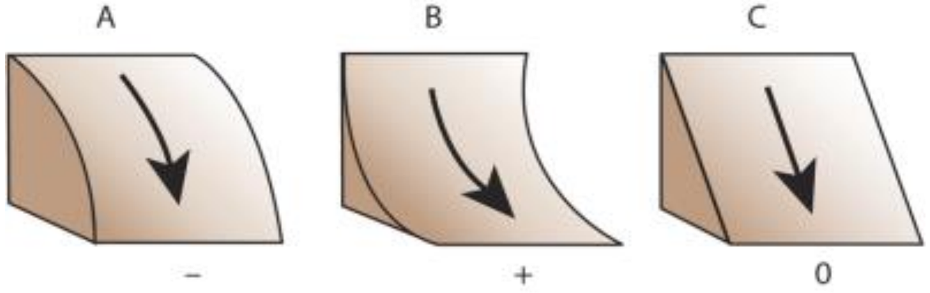
Photogrammetric Engineering & Remote Sensing
Vol. 80, No. 5, May 2014, pp. 439–449.
0099-1112/14/8005–439

doi: 10.14358/PERS.80.5.439

1.) Topographic features



Plan Curvature (across slope)

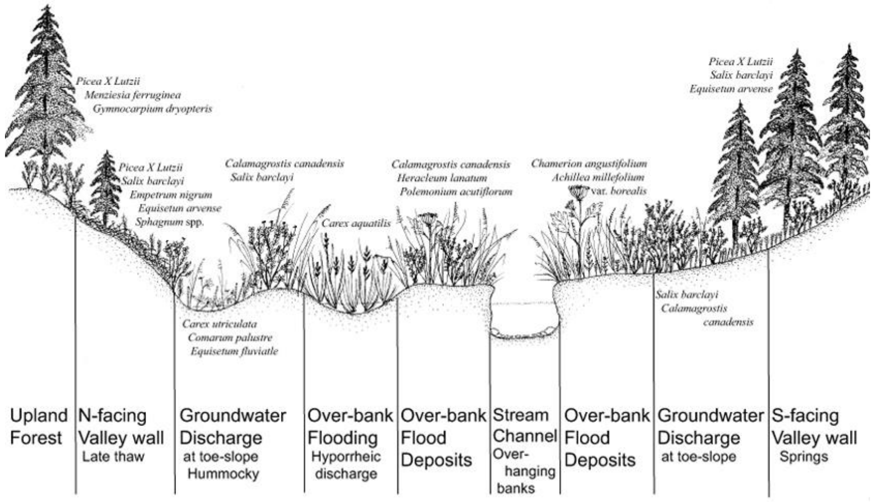


Profile Curvature (along slope)

DEV= (elevation - mean elevation)/standard deviation elevation

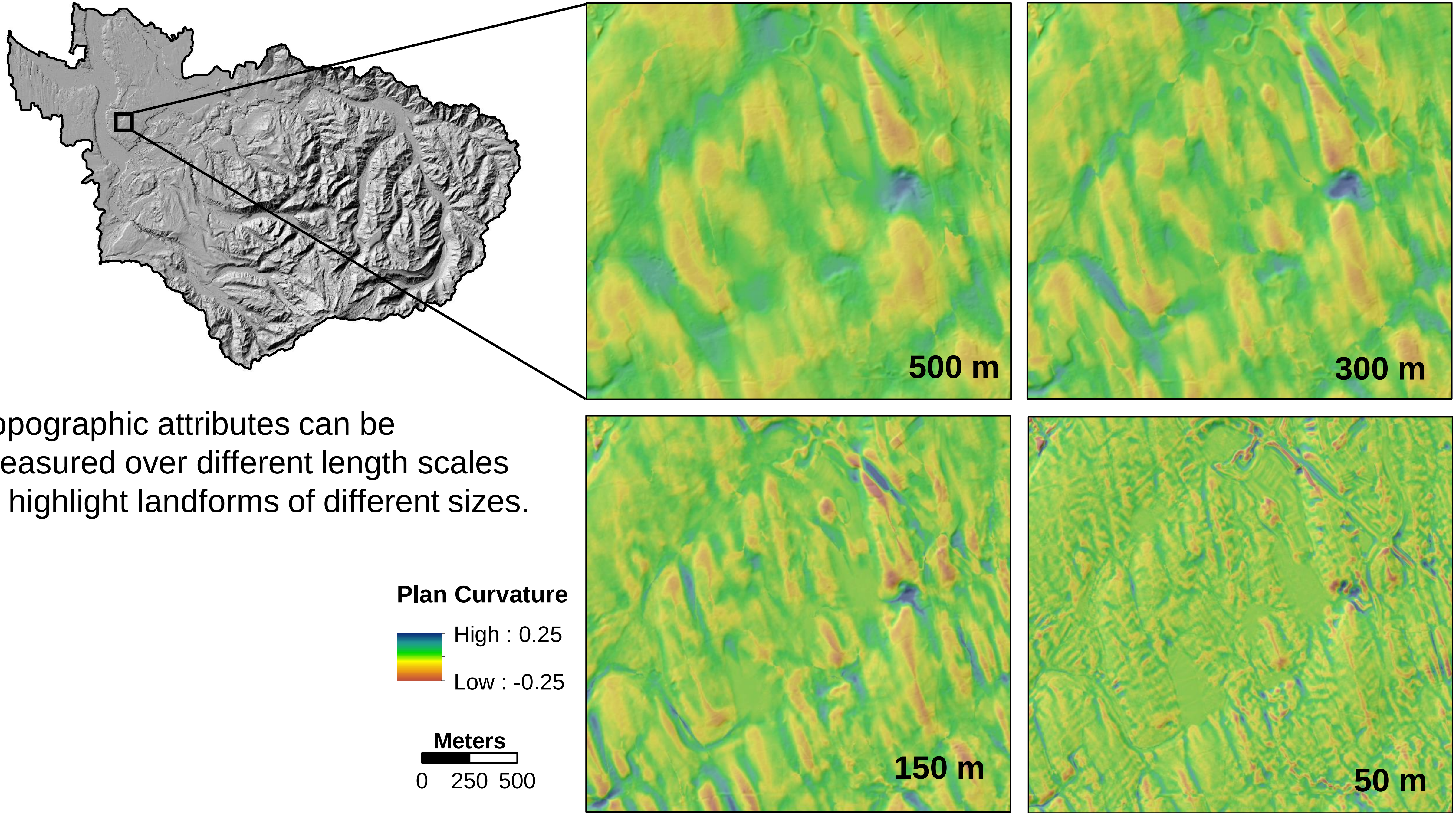
Topographic Indices :

- plan curvature,
- profile curvature,
- gradient (slope)
- DEV

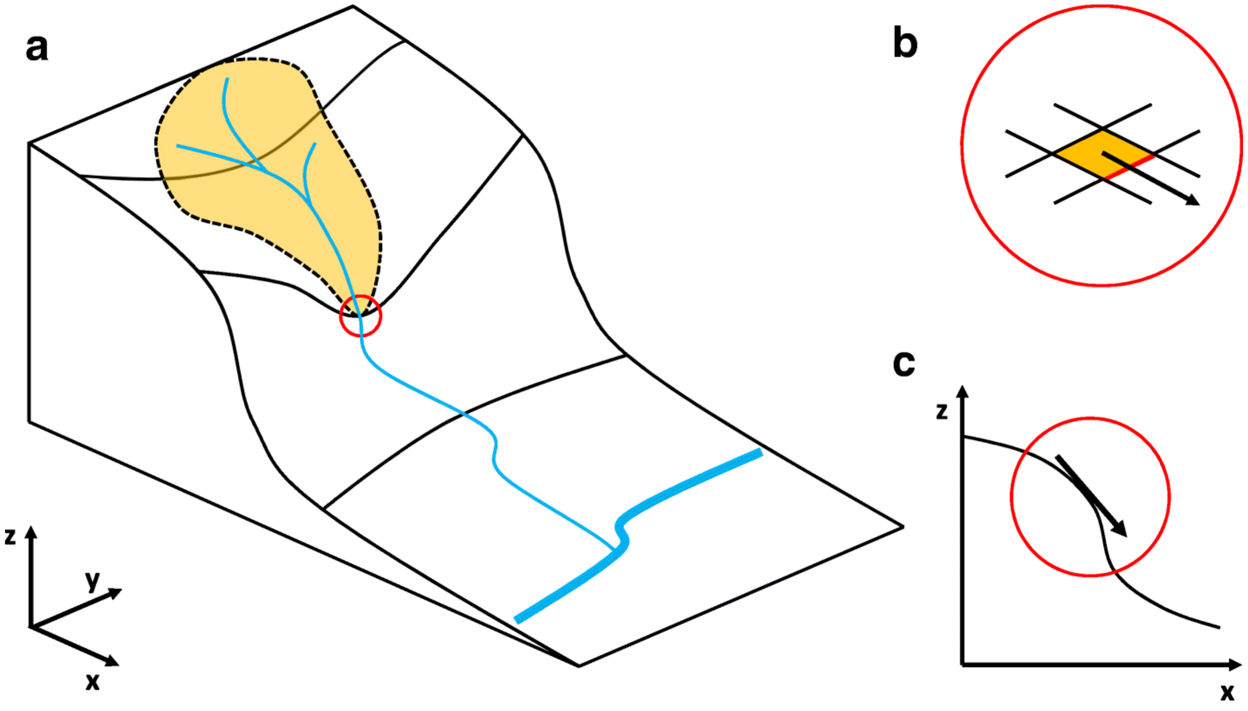


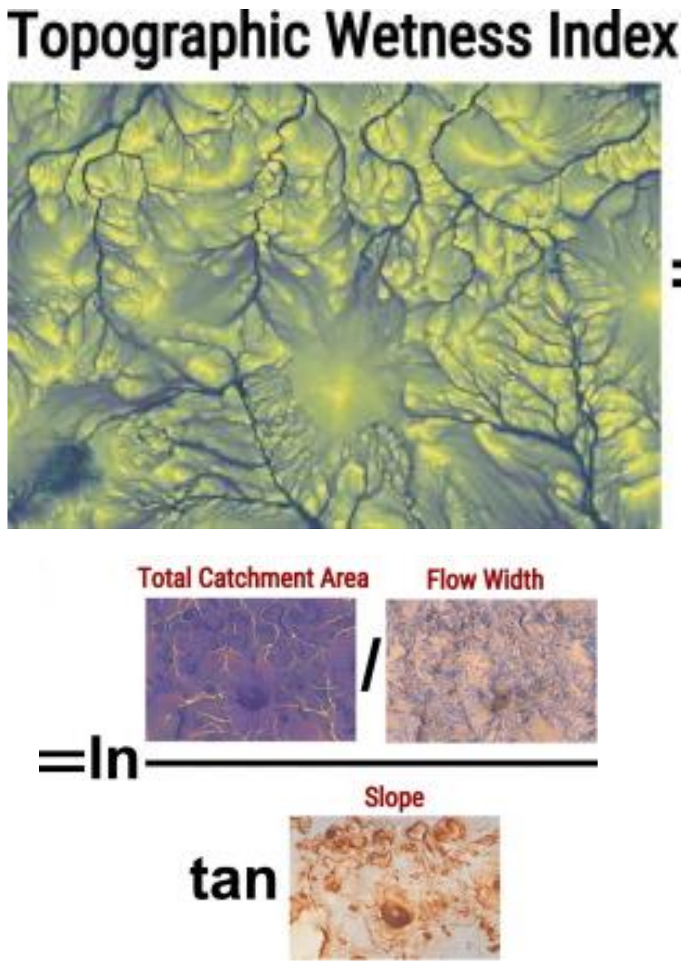
Artwork by Conrad Field

Topographic attributes can be measured over different length scales to highlight landforms of different sizes.



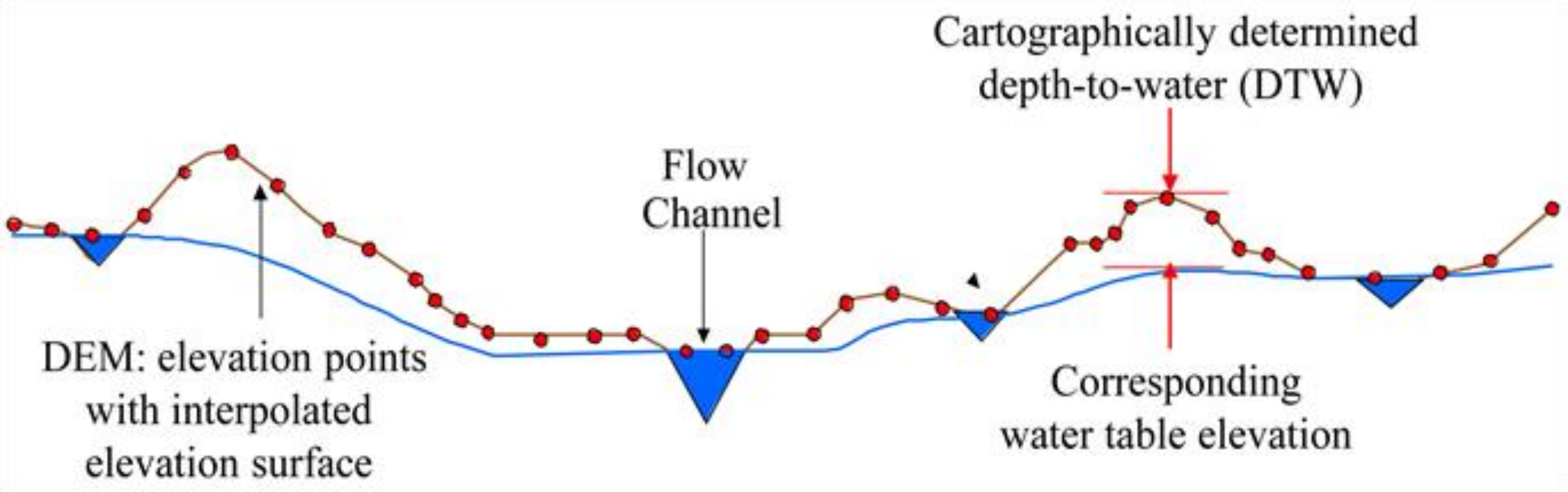
2.) Hydrologic Modelling - TWI

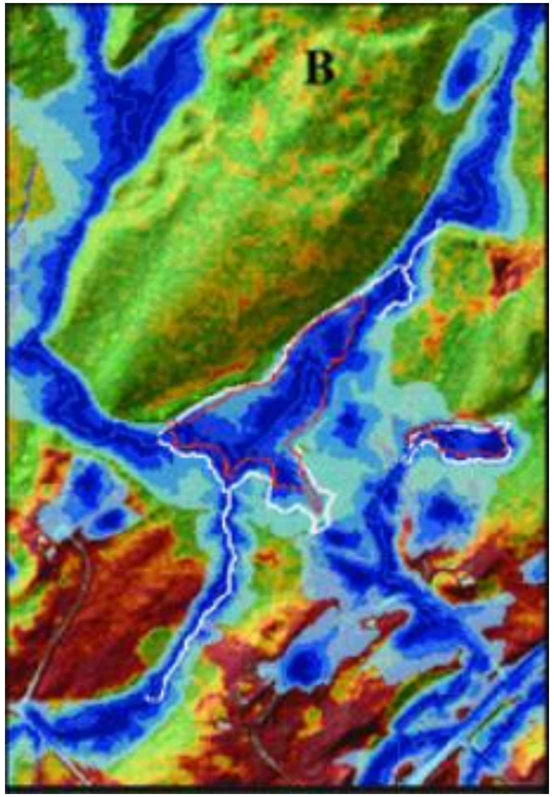




Martin Kopecký, Martin Macek, Jan Wild,
Topographic Wetness Index calculation guidelines based on measured soil moisture and plant species composition,
Science of The Total Environment, Volume 757, 2021,https://doi.org/10.1016/j.scitotenv.2020.143785.

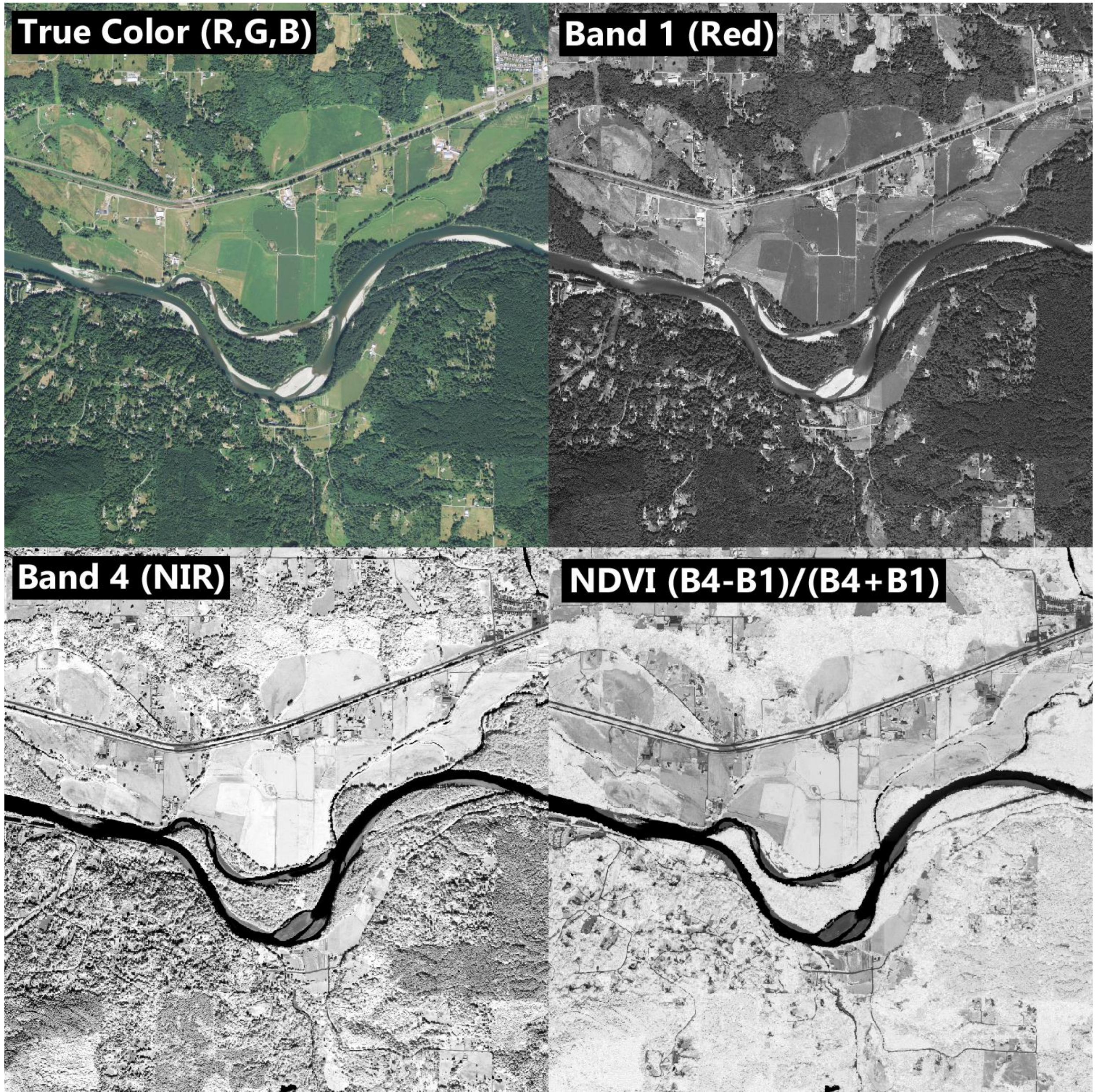
2.) Hydrologic Modelling - Depth-to-water index



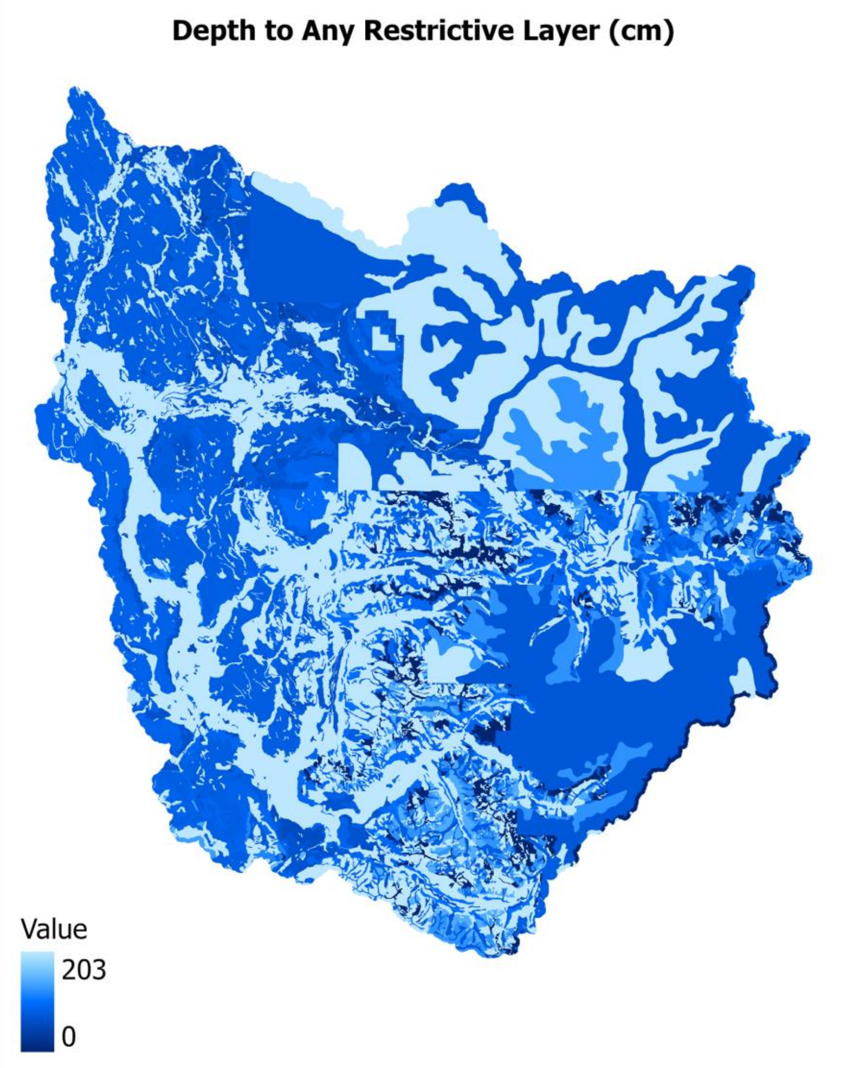


White, Barry & Ogilvie, Jae & Campbell, David & Hiltz, Douglas & Gauthier, Brian & Chisholm, H. & Wen, Hua & Murphy, Paul & Arp, Paul. (2012). Using the Cartographic Depth-to-Water Index to Locate Small Streams and Associated Wet Areas across Landscapes. Canadian Water Resources Journal. 37. 10.4296/cwrj2011-909.

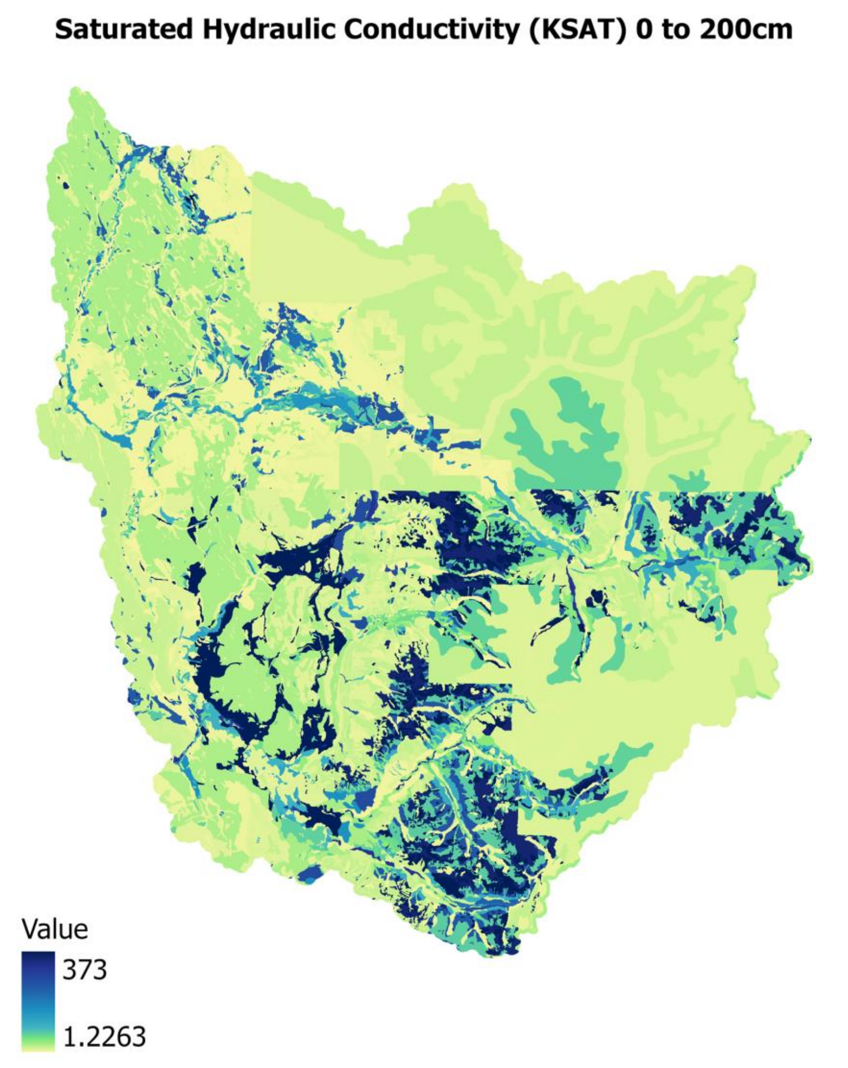
3.) Spectral indices - NDVI



4.) Vector datasets – soils, geology, other wetland inventories



Depth to Any Restrictive Layer (cm)



Saturated Hydraulic Conductivity (KSAT) 0 to 200cm

Collect training and validation data collection

Total points = 2,417

GRTS sample design:

Puyallup

- 1,270 point photo interpreted
- 101 assessed in the field

Mashel:

- 94 points photo interpreted
- 74 assessed in field

WIP Tool sample design:

Coulter Creek:

410 points photo interpreted.

Spent 5 days in the field.

36 assessed in the field.

Hoh watershed:

360 points photo interpreted.

Spent 5 days in the field.

145 assessed in the field

Machine Learning – Random forest model

slides credit: Keenan Ganz

- Random forests are built from decision trees
- Place **observations** into **classes** by making binary decisions on their **features**

- In this study:
  - Observations: Training points in the study area
  - Classes: wetland/upland
  - Features: data from input rasters

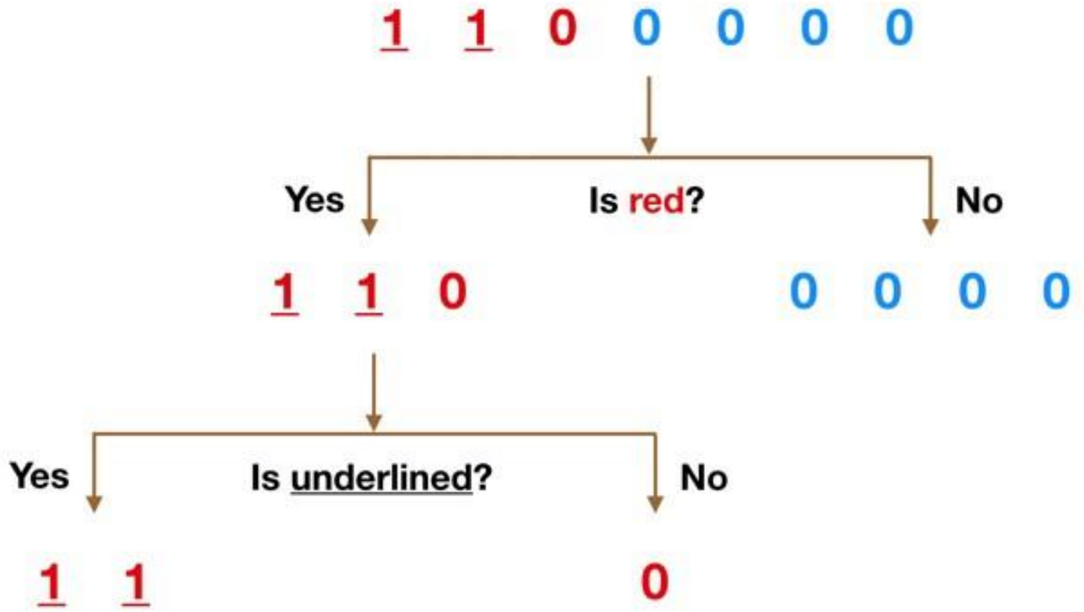
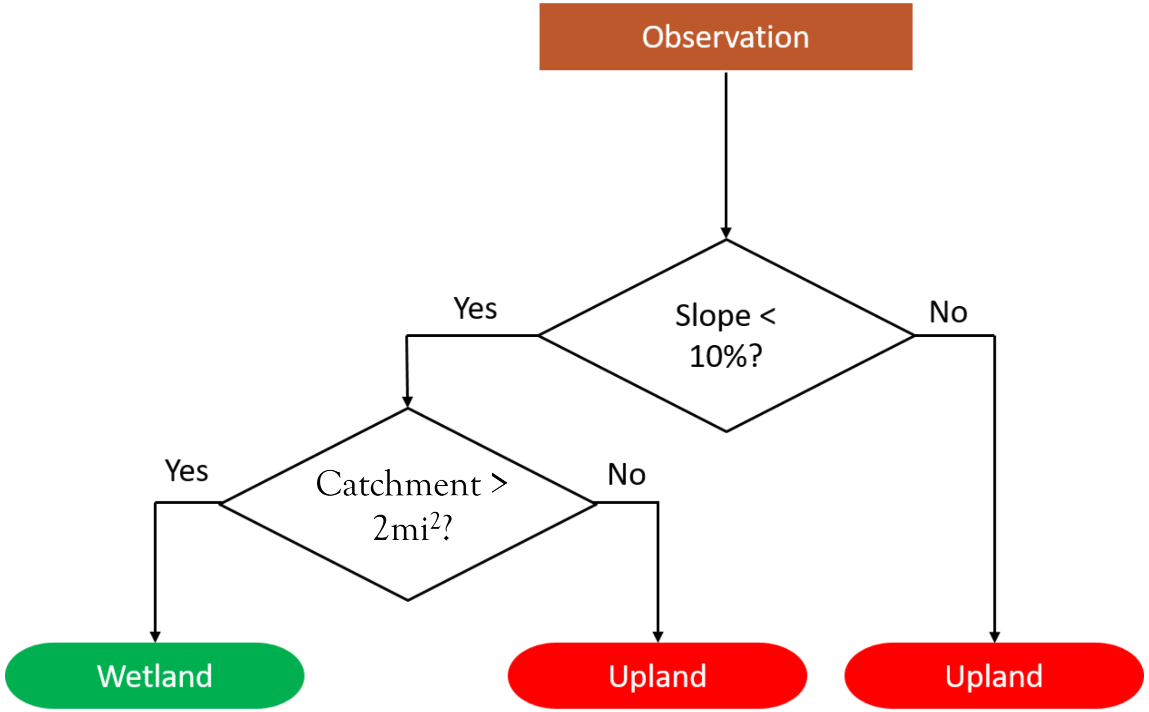


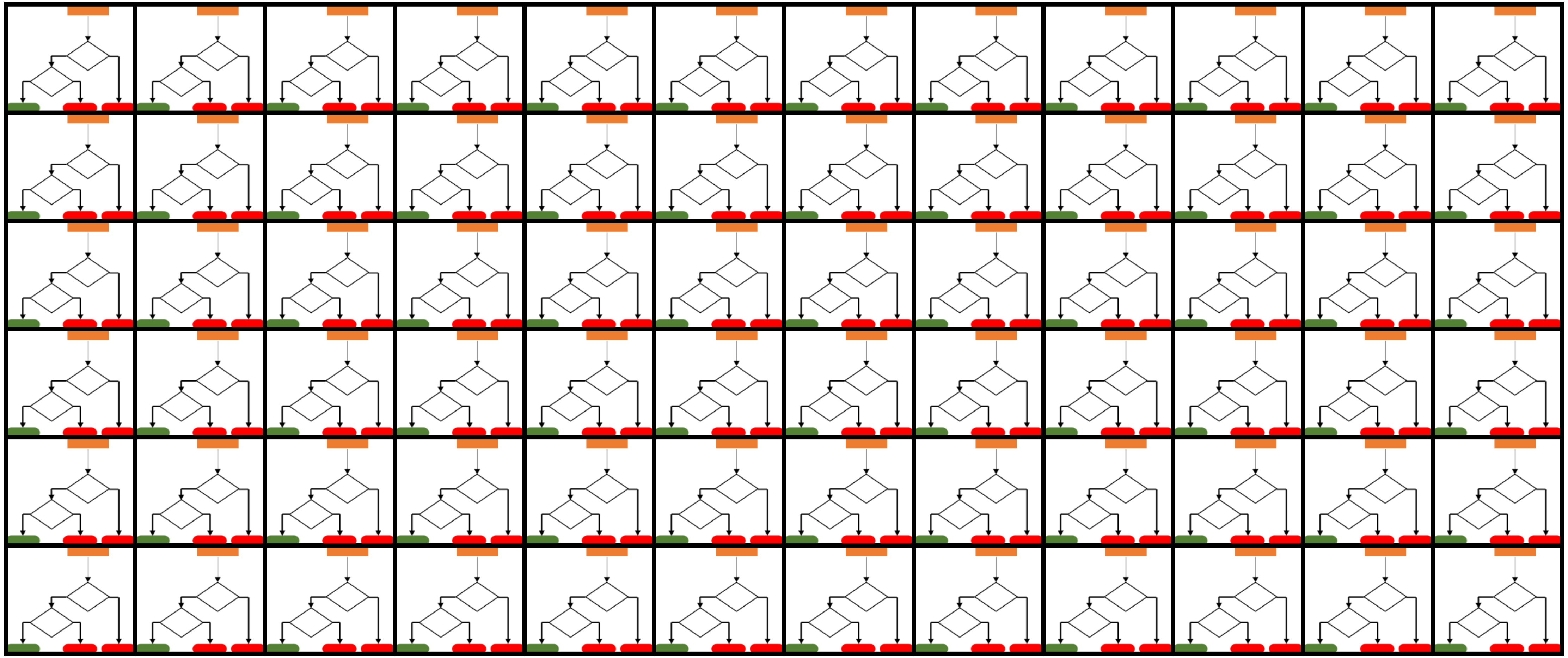
Image: Tony Yiu

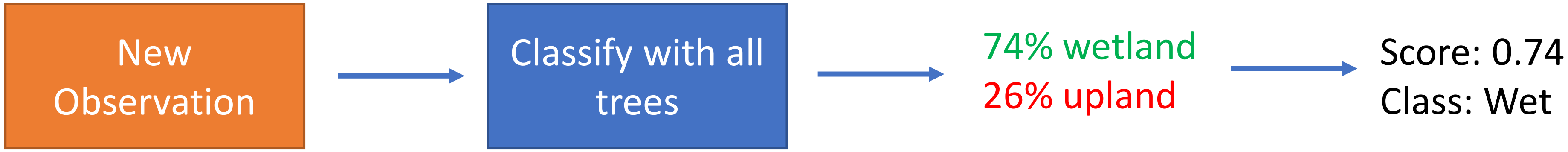
Random forests are built from decision trees

- Random forests generate 100s-1000s of decision trees, built on unique subsets of observations and features in the training data

- Classifications are predicted by taking a vote of all trees in the model







New Observation

Classify with all trees

74% wetland
26% upland

Score: 0.74
Class: Wet

RESULTS

Results – Wetland probability for Puyallup watershed

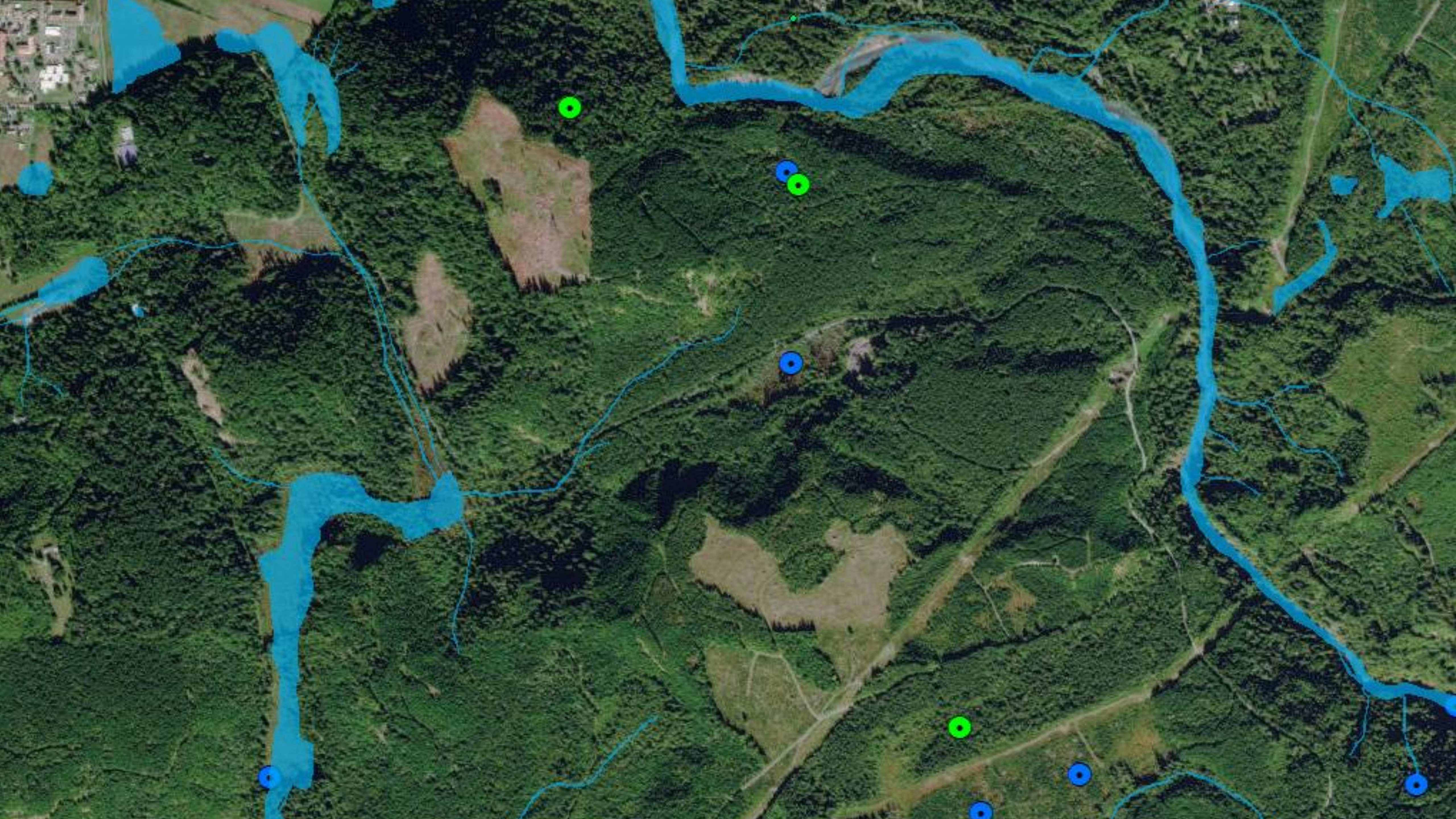
- Model outputs = Probability raster, 0 – 1 likelihood
- Using a cutoff of 0.5= 4 x more wetland area in forested areas than the NWI

NWI
Overall accuracy = 88.1%

Error of Commission = 2.1%
Error of Omission = 41.8%

WIP model
Overall accuracy = 96.6%

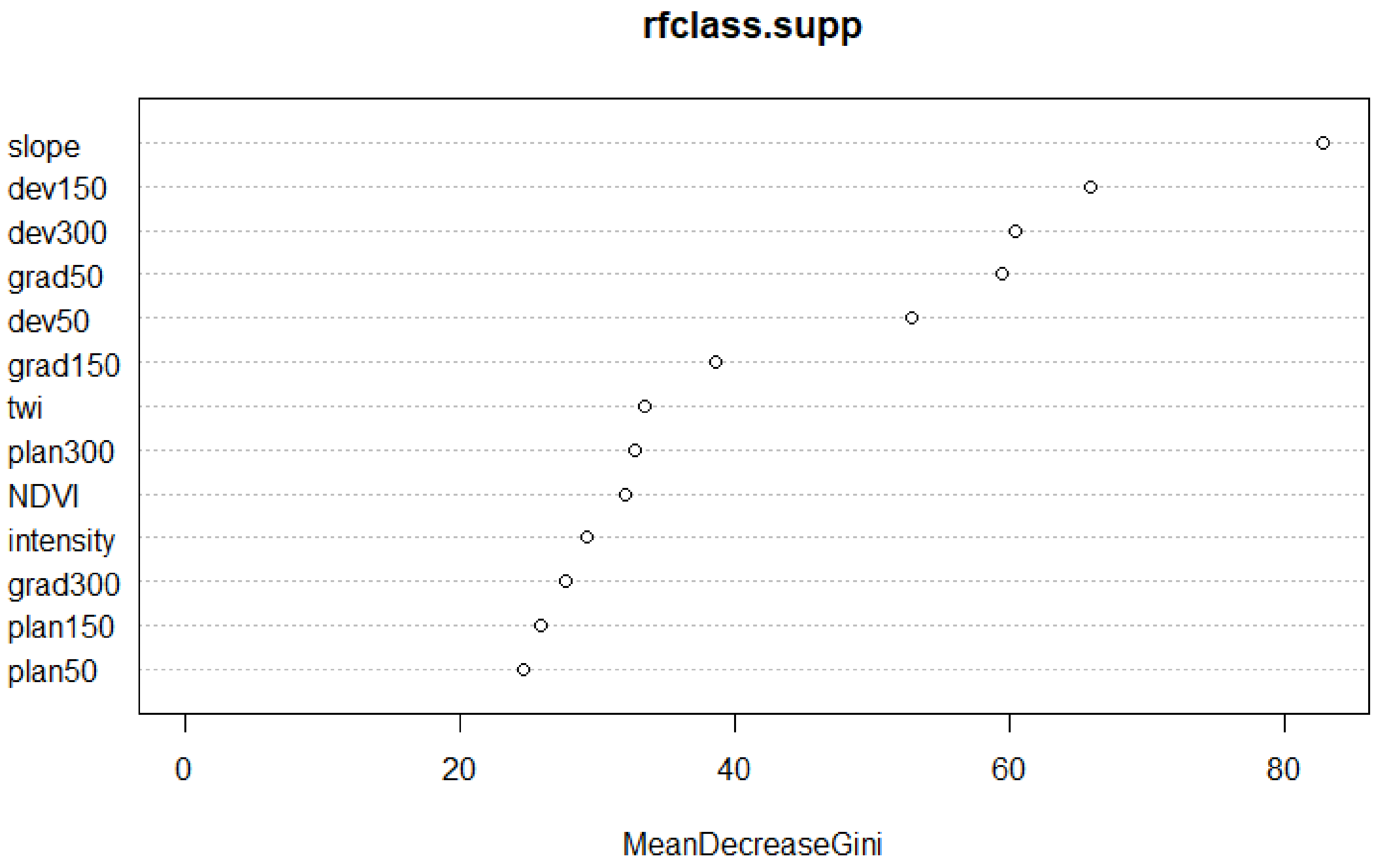
Error of Commission = 4.3%
Error of Omission = 8.0%




High : 1
Low : 0

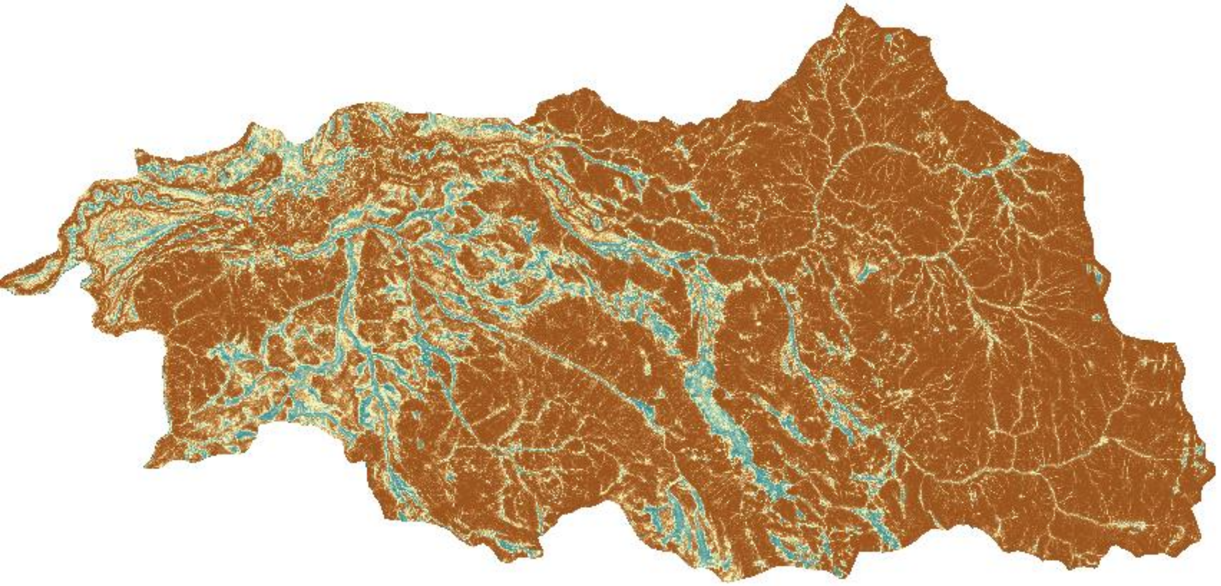
Random Forest Model Variable Importance

Feature importance for full random forest model for Puyallup. The mean decrease in Gini coefficient is a measure of how each variable contributes to the homogeneity of the nodes and leaves in the resulting random forest. Variables at the top contributed the most definition in the random forest model.

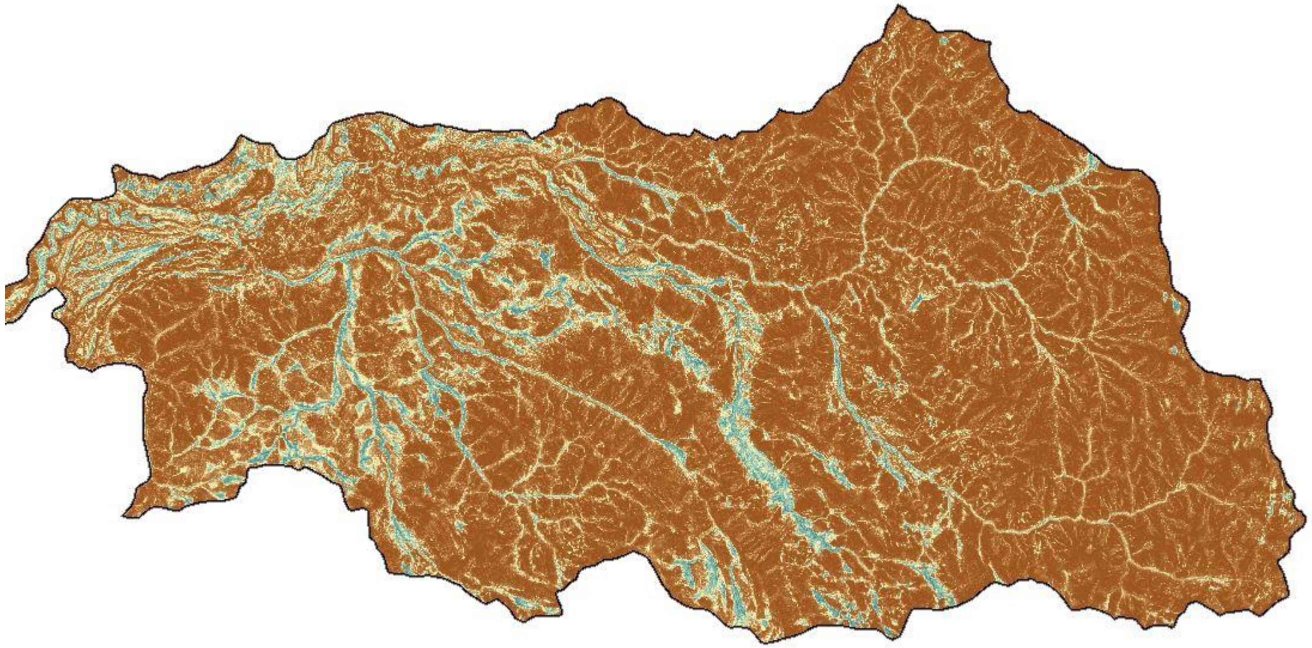


A model trained on a similar area can be transferred to that area

Mashel watershed



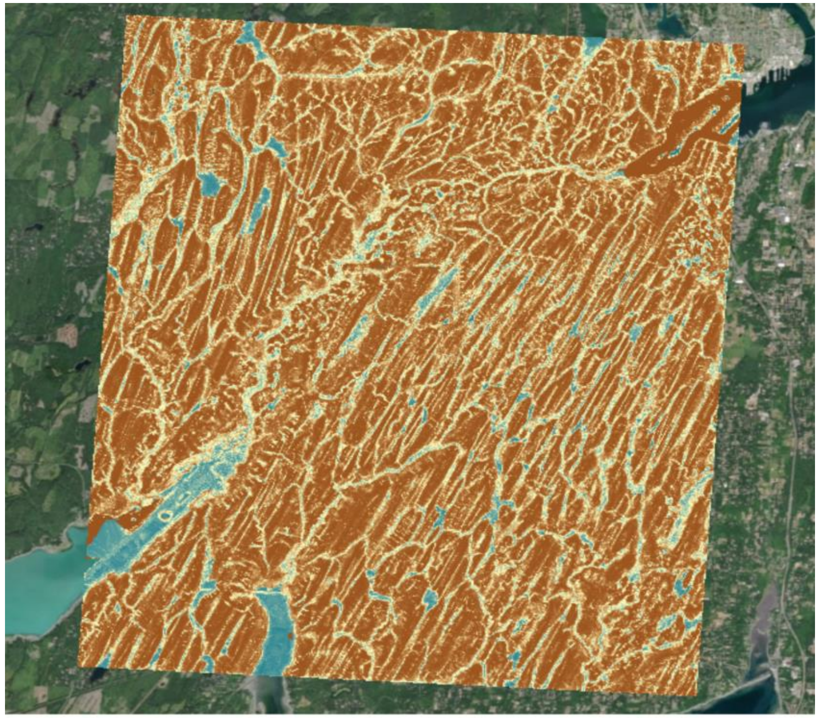
Overall accuracy = 97%
Error of omission = 16%



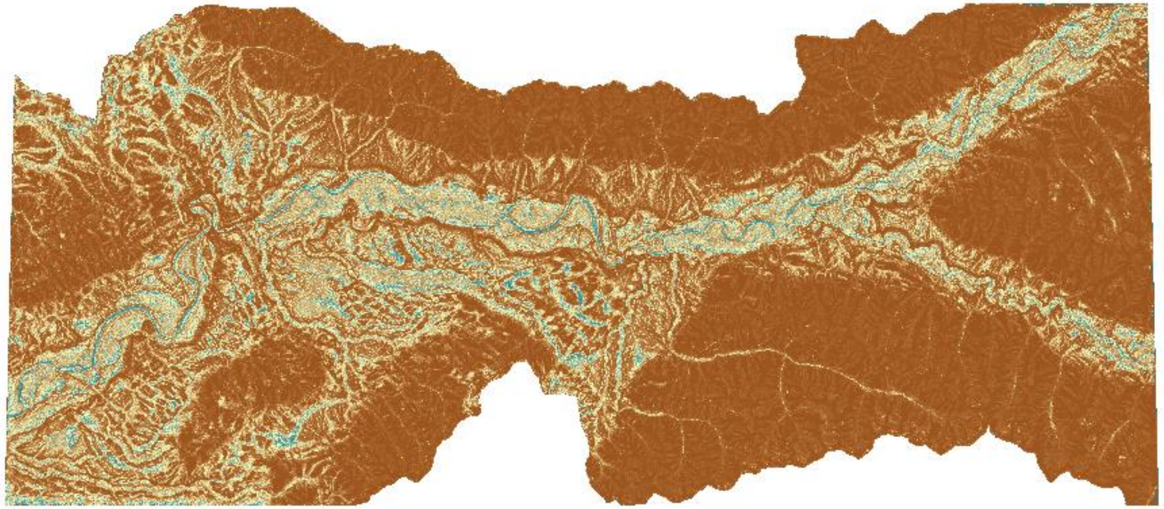
Overall accuracy = 96%
Error of omission = 21%

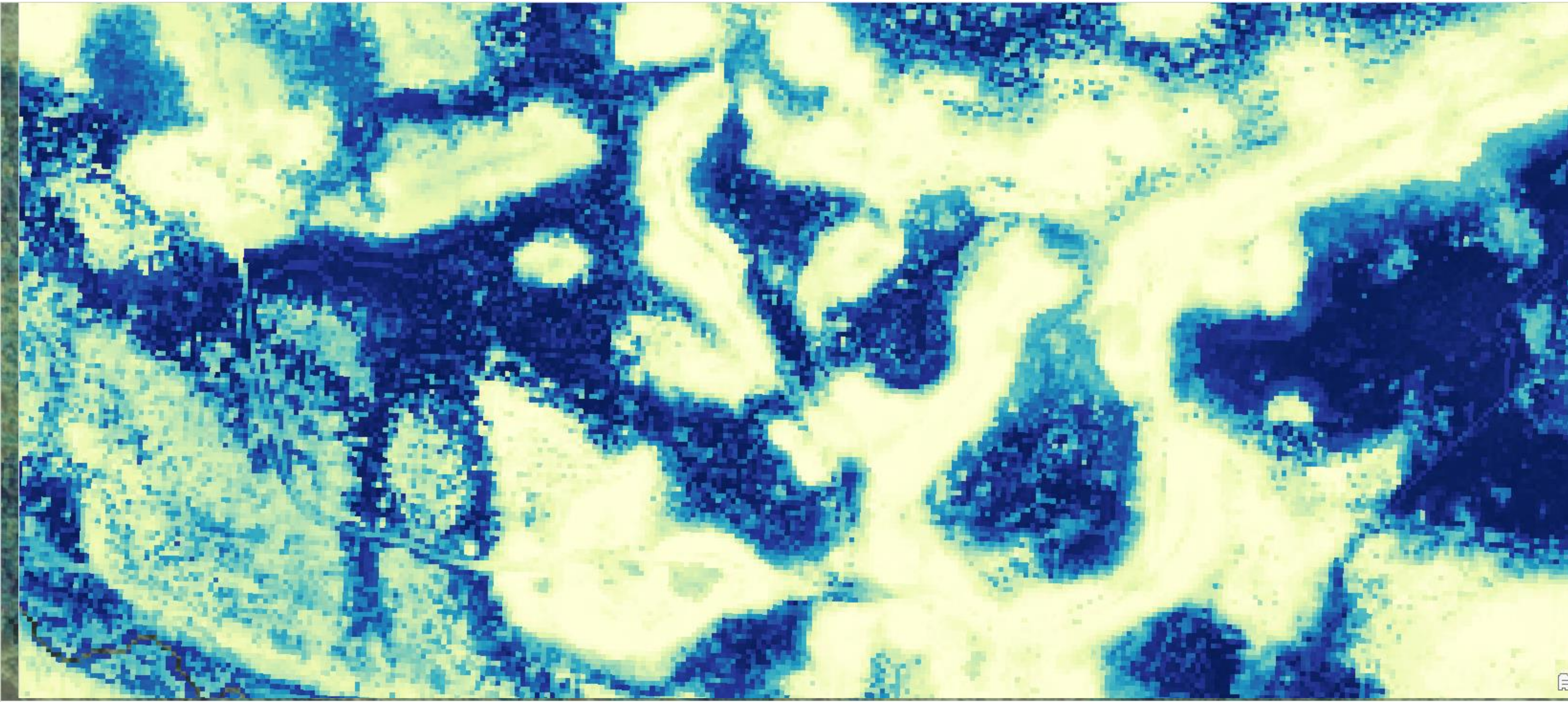
However, need to build a new model for areas that are in a different ecoregion – Easy to do!

Coulter Creek – Kitsap peninsula



Hoh watershed





Conclusion

- The WIP tool identifies wetlands missed in existing wetland inventories
- These may be wetlands that are hard to identify in aerial imagery alone.
- The model can be improved as new input data layers are identified as important.
- Can be used for improving sampling efficiency
- Can be used to screen for potential wetlands – can lower the cutoff or raise the cutoff.
- WIP model performs better when field data is used, but works very well with NWI training data (available everywhere).

Limitations of the WIP Tool:

The WIP tool provides an improvement on identifying wetland locations in forested areas, but does not delineate wetland borders or classify wetland types. For any policy or management application, the WIP tool is best used as an initial screening for follow-up on the ground.

There are several limitations of the WIP tool:
1.) We did not use a jurisdictional wetland definition.

2.) While in theory, the WIP tool should effectively map wetlands in Eastern Washington, none of our study areas for this project were located in Eastern Washington.

Limitations of the WIP Tool (Cont…) :

4.) The WIP tool may not provide useful results for slope wetlands and these wetlands will likely be missed in any WIP tool product. We did not have adequate training data locations of slope wetlands to train our model and therefore we could not test out the effectiveness of mapping slope wetlands using the WIP tool.

5.) The WIP tool is based on topographic features and surface water flow models. It does not account for well-drained soils. Certain areas may identify strongly as wetlands, but in fact be false positives due to underlying geology and soil types.

Limitations of the WIP Tool (Cont…) :

6.) The WIP tool was created primarily for forested wetlands. It may be useful for other non-forested areas, but this was not the focus of this research, and therefore it has not been assessed.

7.) The WIP tool may not produce useful results for areas with constructed human modification of water flows (i.e. drains, ditches) as these are not mapped as part of the lidar-derived hydrologic flow models used as inputs to the WIP tool.

Updates : Continuing to improve WIP tool

Increasing data in the Hoh watershed

- 800 additional points (1/4 visited in the field)
- Exploring additional soils layers
- Extending to the Colville watershed (eastside).
- Running for the Snohomish county

Some of the uses of the tool:

- Mapping wetlands on forestlands – screening tool
- Used for sampling design for forested wetlands.
- Identifying wetlands to consider how they might mitigate drought and stream permanence by recharging groundwater and storing water in hotter summer months. (Tulalip tribe)
- Currently using the WIP tool for a NASA study to map below ground carbon sequestration on forestlands along the wet to dry gradient, and specifically in forested wetlands.

Questions

Meghan Halabisky, halabisk@uw.edu